Role of hydration and intramolecular interactions in the helix-coil transition and helix-helix assembly in a deca-alanine peptide

Dheeraj S. Tomar,[1] Valéry Weber,[2] B. Montgomery Pettitt,[3] and D. Asthagiri[3, 4, *]

[1]*Department of Chemical and Biomolecular Engineering,*
Johns Hopkins University, Baltimore, MD 21218
[2]*IBM Research, Zurich, CH-8803 Rüschlikon, Switzerland*
[3]*Sealy Center for Structural Biology and Molecular Biophysics,*
Department of Biochemistry and Molecular Biology,
University of Texas Medical Branch, Galveston, TX 77555
[4]*Department of Chemical and Biomolecular Engineering, Rice University, Houston, TX 77005*
(Dated: August 25, 2015)

For a model deca-alanine peptide the cavity (ideal hydrophobic) contribution to hydration favors the helix state in the coil-to-helix transition and the paired helix bundle in the assembly of two helices. The energetic contributions of attractive protein-solvent interactions are separated into a short-range part arising from interactions with solvent in the first hydration shell and the remaining long-range part. In the helix-coil transition, short-range attractive protein-solvent interactions outweigh hydrophobic hydration and favor the unfolded coil states. Analysis of enthalpic effects shows that it is the favorable hydration of the peptide backbone that favors the unfolded state. Protein intramolecular interactions favor the helix state and are decisive in folding. In the pairing of two helices, the cavity contribution outweighs short-range attractive protein-water interactions. However, long-range, protein-solvent attractive interactions can either enhance or reverse this trend depending on the mutual orientation of the helices. In helix-helix assembly, change in enthalpy arising from change in attractive protein-solvent interactions favors disassembly. In helix pairing as well, favorable protein intramolecular interactions are found to be as important as hydration effects.

Keywords: protein folding, driving force, protein hydration free energy, molecular dynamics

Helices have been termed the "hydrogen atoms of biomolecular complexity" [1]. In proteins the α-helix is a common structural motif and understanding the formation of α-helices occupies a pre-eminent place in efforts to understand protein folding. Using computer simulations and a new approach to free energy calculations [2, 3], here we revisit this classic problem and study two transitions in a model deca-alanine peptide. Mirroring the primary-to-secondary and secondary-to-tertiary transitions in protein folding, we study, respectively, the coil-to-helix transition and the pairing of helices to form a helix dimer. Our principal focus is to explicate the hydrophobic and hydrophilic hydration contributions in these transitions at a level that has hitherto been possible only for small molecular solutes.

There are several reasons to re-examine the fundamental premises about hydration thermodynamics of proteins and the forces driving protein folding. First, recent experiments and simulations challenge the conventional view that hydrophobic interactions drive protein folding. These studies show that the all-backbone polyglycine, and analogous archetypes of intrinsically disordered peptides, can undergo a collapse transition in water [4–8]. Second, recent computer simulations [9–12] reveal important physical and conceptual limitations in the group-additive approach that has been the backbone of approaches to understand protein hydration thermodynamics in experiments (cf. Ref. 13 and 14). Indeed the prevailing views of protein hydration thermodynamics [15–17] trace back to this group-additive reasoning. Lastly, while simulations can in principle provide a detailed molecular thermodynamics understanding of hydration, with some exceptions [18, 19], such studies are scarce for realistic proteins and polypeptides. The availability of an approach that alleviates this situation presents an opportunity to revisit and potentially re-appreciate a classic problem in protein folding.

Earlier studies based on continuum solvent or lattice models have come to differing conclusions about solvent effects in the coil-to-helix transition. Some have suggested that hydrophobicity drives the transition [20, 21], while others have emphasized the role of favorable electrostatics [22]. Experiments suggest that helix extension is enthalpically driven [23], as has also been found in computer simulations interpreted within the Zimm-Bragg or Lifson-Roig formalisms (for example, see Refs. 24–26). Interestingly, both experiments [27] and computer simulations [24–26] show a negative heat capacity upon unfolding, the opposite of what is observed in unfolding of globular proteins [28]. The negative heat capacity has been interpreted as arising due to the favorable hydration of the backbone upon unfolding [27], but a molecular scale description of this signature remained to be sought.

The approach we have developed is based on a quasichemical organization [29–31] of the potential distribution theorem [32]. Using this approach, we are able to interrogate the hydration thermodynamics of proteins [2, 3] at a level of resolution that is comparable to those

for small molecular solutes. Our study on cytochrome C helped reveal limitations of continuum models of hydration [3]. Subsequent studies have helped illuminate conceptual and physical limitations in the group-additive description of the hydration thermodynamics of the peptide backbone[11] and of a hydrophobic side-chain in the context of model peptides [12]. On the basis of these new developments, and for the reasons noted above, here we study the hydration thermodynamics in the coil-to-helix and helix-helix assembly in a deca-alanine peptide.

We find that hydrophobic interaction is not consequential in the formation of the helix, but it does play an important role in helix-dimerization. But hydrophilic hydration is found to play a nontrivial role in the coil-to-helix transition and helix-dimerization. Indeed hydrophilic hydration can drive the unfolding of the helix and disassembly of the helix dimer. Throughout backbone hydration is found to be the most important component of the overall hydrophilic hydration. For the systems studied here, hydrophilic hydration and protein intramolecular interactions are as important as, if not more important than, hydrophobic effects.

I. THEORY

The excess chemical potential, μ^{ex}, is that part of the Gibbs free energy that arises from intermolecular interactions and of first interest in understanding the solubility of a solute in a solvent. Here μ^{ex} is defined relative to the ideal gas at the same density and temperature. To calculate μ^{ex}, we regularize the statistical problem of calculating it from the potential distribution theorem $\beta\mu^{\mathrm{ex}} = \ln\langle e^{\beta\varepsilon}\rangle$ [30, 32], where the averaging $\langle\ldots\rangle$ is over the solute-solvent binding energy (ε) distribution $P(\varepsilon)$. As usual $\beta = 1/k_{\mathrm{B}}T$, with T the temperature and k_{B} the Boltzmann constant.

We introduce an auxiliary field $\phi(\lambda; r)$ that moves the solvent away from the solute to a distance λ. The distance between the center of the field and the solvent molecule is r. For $r > \lambda$, $\phi = 0$. Since the solvent interface is pushed away from the solute, the solute-solvent interaction is tempered and the conditional distribution $P(\varepsilon|\phi)$ is better behaved than $P(\varepsilon)$ [2, 3, 11]. In practice, we adjust the range λ such that $P(\varepsilon|\phi)$ is a Gaussian. With the introduction of the field, we have[2, 3, 11]

$$\beta\mu^{\mathrm{ex}} = \underbrace{\ln x_0[\phi]}_{\text{local chemistry}} + \underbrace{\beta\mu^{\mathrm{ex}}[P(\varepsilon|\phi)]}_{\text{long-range}} \underbrace{-\ln p_0[\phi]}_{\text{packing}}, \quad (1)$$

where $-k_{\mathrm{B}}T\ln x_0[\phi(\lambda)]$ is the work done to apply the field in the presence of the solute, $-k_{\mathrm{B}}T\ln p_0[\phi(\lambda)]$ is the corresponding quantity in the absence of the solute, and $\beta\mu^{\mathrm{ex}}[P(\varepsilon|\phi)]$ is the contribution to the interaction free energy in the presence of the field. Fig. 1 provides a schematic description of Eq. 1.

We apply the field about each heavy atom to carve out the molecular shape in the liquid (Fig. 1). For convenience we use the same value of λ for all the heavy
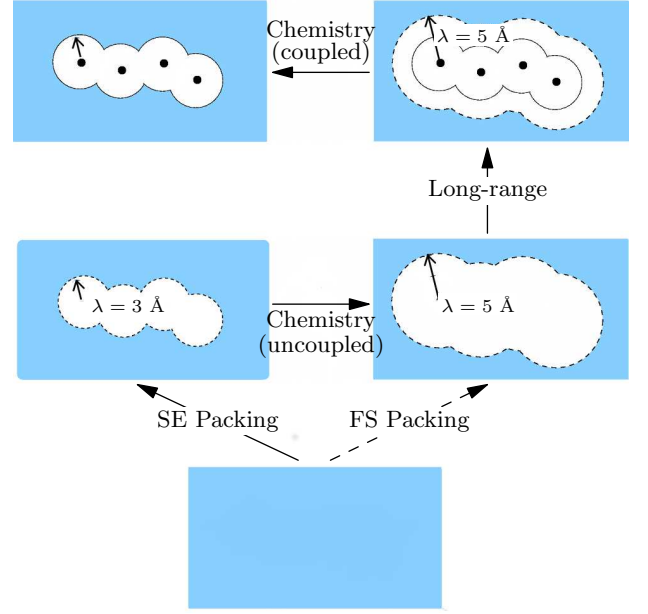

FIG. 1. Quasichemical organization of the excess chemical potential. The $\lambda = 3$ Å envelope defines the solvent excluded (SE) volume and the $\lambda = 5$ Å defines the envelope extending to the first hydration shell (FS). For chemistry coupled (uncoupled), the solute-solvent interaction is present (absent). In Eq. 1 we follow FS-packing to the hydrated solute; in Eq. 2 we follow SE-packing. Figure adapted from Ref. 11 with permission from Elsevier.

atoms. For $\lambda \geq 5$ Å, $P(\varepsilon|\phi)$ is well-described by a Gaussian. We thus define $\lambda_{\mathrm{G}} = 5$ Å.

The solute also excludes a volume to the solvent and by definition there are no short-range attractive interactions between the solute and the solvent in this domain. We find that $\ln x_0 \approx 0$ for $\lambda \leq 3.0$ Å, irrespective of the conformation of the peptide. This suggests that the space enclosed by $\lambda_{\mathrm{SE}} = 3.0$ Å is excluded to the solvent and thereby provides a natural definition of molecular extent of the cavity to be used in discussions of the cavity (ideal hydrophobic) contribution. Interestingly, the range between 3 Å to 5 Å corresponds to the first hydration shell for a methyl carbon[33] and is a conservative description of the first hydration shell of groups containing nitrogen and oxygen heavy atoms. For simplicity

Using the λ_{SE} and λ_{G}, we rearrange Eq. 1 as

$$\beta\mu^{\mathrm{ex}} = \underbrace{\ln \frac{x_0[\phi(\lambda_{\mathrm{G}})]}{p_0[\phi(\lambda_{\mathrm{G}})]/p_0[\phi(\lambda_{\mathrm{SE}})]}}_{\text{revised chemistry}} + \underbrace{\beta\mu^{\mathrm{ex}}[P(\varepsilon|\phi(\lambda_{G}))]}_{\text{long-range}}$$
$$\underbrace{-\ln p_0[\phi(\lambda_{\mathrm{SE}})]}_{\text{SE packing}} . \quad (2)$$

Note that SE packing and the revised chemistry plus long-range contribution is uniquely defined for the given forcefield. Physically, the revised chemistry contribution measures the free energy contribution from solute inter-

action with solvent in the first shell relative to a non-interacting (uncoupled) solute (Fig 1).

The revised chemistry and long-range contributions describe the role of short-range and long-range attractive protein-solvent interactions in the thermodynamics of hydration. These two components of the hydrophilic contribution occur at different energy (and length) scales. But within commonly used continuum models of electrostatic interactions this distinction is necessarily lost and both the short-range and long-range interactions are treated as part of long-ranged interactions (e.g. see [17]). From the perspective of such continuum models, our definition of long-range interaction is more conservative.

The SE packing contribution measures the hydrophobic hydration of an ideal hydrophobe [34, 35]. In theoretical discussions of hydrophobic effects, a cavity with a hard-wall interaction in water is often considered. The packing contribution in our calculation uses a soft-cavity. This soft-cavity packing estimate is always a lower-bound to the hard-cavity estimate and can be easily corrected to give the latter [2]. We do not pursue those corrections here and instead use the soft-cavity packing result as a measure of primitive hydrophobic effects. We refer the reader to published papers [2, 3, 11] for more extensive details about the approach.

A. Helix-Helix PMF

The above development carries over to the calculation of $W(r)$, the potential of mean force (PMF) to bring two helices a distance r apart, where r is the separation between the helix axis with the axis parallel to each other. The PMF

$$W(r) = W_{solv}(r) + \Delta U(r), \qquad (3)$$

where W_{solv} is the solvent (or indirect) contribution [33] and ΔU is the contribution from direct protein-protein interactions. $W_{solv}(r) = \mu^{\mathrm{ex}}_{\mathrm{dimer}}(r) - 2\mu^{\mathrm{ex}}$, where $\mu^{\mathrm{ex}}_{\mathrm{dimer}}(r)$ is the hydration free energy of the pair of helices (for a given separation and orientation) and μ^{ex} is the hydration free energy of a monomer helix.

B. Entropic and enthalpic contributions

Ignoring the pressure-volume correction and contribution due to a finite isothermal compressibility of water, the enthalpy of hydration, h^{ex}, is given by

$$h^{\mathrm{ex}} = E_{sw} + E_{reorg} \qquad (4)$$

where E_{sw} is the average peptide-solvent interaction energy and E_{reorg} is the (average) water reorganization energy. Additionally ignoring the contribution due to a finite thermal expansivity of water, the entropy of hydration is given by

$$Ts^{\mathrm{ex}} = h^{\mathrm{ex}} - \mu^{\mathrm{ex}}. \qquad (5)$$

For calculating E_{reorg}, we adapted the hydration-shell-wise calculation described earlier [12, 33] (cf. Sec. S.III).

II. METHODS

The regularization follows previous work [11]. The deca-alanine peptide was modeled with an acetylated (ACE) N-terminus and n-methyl-amide (NME) capped C-terminus. The extended β-conformation ($\phi, \psi = -154 \pm 12, 149 \pm 9$) was aligned such that the end-to-end vector lay along the diagonal of the simulation cell. We label this coil state as C_0. The helix was aligned with the helix-axis along the x-axis of the cell. The initial structures were energy minimized with weak restraints on the heavy atoms to relieve any strain in the structure. The peptides were solvated in 3500 TIP3P [36, 37] water molecules. Version c31 of the CHARMM [38] forcefield with correction(cmap) terms for dihedral angles [39], was used for the peptides.

We sampled unfolded states using the adaptive-bias force (ABF[40, 41]) approach which also additionally provided the free energy of unfolding the polypeptide in vacuum. From the ABF trajectory, we sampled nine structures with end-to-end distances between terminal carbon atoms ranging between 28 Å and 36 Å in increments of 1 Å. We label the coil states from this unfolding simulation $\{C_1, \ldots, C_9\}$. The ϕ, ψ for these unfolded states predominantly populate β and PPII regions of the Ramachandran plot. (Note that, in the strict sense of coil-to-helix transition, the extended β is one extreme member in the ensemble of coil states.) In all the hydration free energy calculations, the structures were held rigid.

The same set-up was used to investigate helix pairing. Additionally, we consider two relative orientations of the helix dipoles, parallel and antiparallel. (Note that the helix dipoles will be antiparallel in the simplest helix-turn-helix motif.) These arrangements help illuminate the role of long-range protein-solvent interactions in helix-helix complexation. We note that in nature helices almost never align perfectly parallel or antiparallel [42], but the hydration effects that are of principal interest here are expected to be insensitive to minor distortions in the relative orientation.

III. RESULTS AND DISCUSSION

A. Coil-to-helix transition

Figure 2 collects the results of the hydration free energies of the helix and $\{C_0, \ldots, C_9\}$ coil states; it is clear that the coil states are better hydrated than the helix. In Table I we present the free energy components of the helix state, the least favorably hydrated coil state (C_0), and the most favorably hydrated coil state (C_7).

Before we discuss our results, we first compare our calculations with existing hydration free energy data. As

TABLE I. Components of the hydration free energy for the helix and the least favorably (C_0) and most favorably (C_7) hydrated coil states. For the helix and C_0 states, results with partial charges turned off (indicated by $\mathbf{Q} = \mathbf{0}$) are also noted. R_g is the radius of gyration (relative to the center of mass) and R_c is the end-to-end distance between terminal carbon atoms in Å. SASA is the solvent accessible surface area in Å^2. All energy values are in kcal/mol. Standard error of the mean is given in parenthesis at the 2σ level.

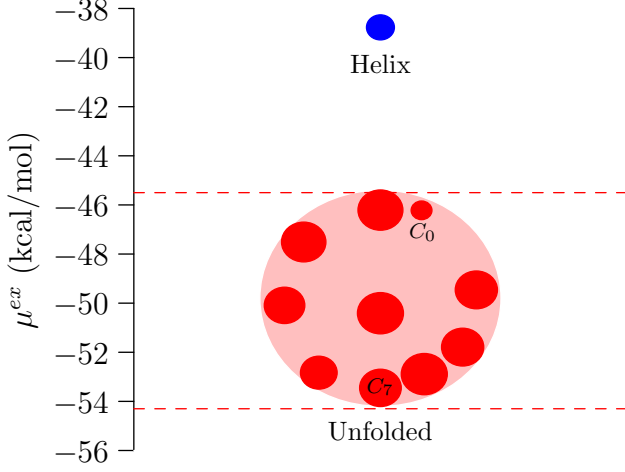| Conformation	R_g	R_c	SASA	SE	Revised Chemistry	Long-range	μ^{ex}	E_{reorg}	E_{sw}	Ts^{ex}
Helix	5.3	16.3	876.0	44.4 (0.4)	−51.6 (0.2)	−31.6 (0.2)	−38.8 (0.5)	75 (6)	−150.9 (0.6)	−37.1
Helix ($\mathbf{Q} = \mathbf{0}$)					−22.4 (0.3)	−18.1 (0.03)	3.9 (0.5)	42 (4)	−59.7 (0.6)	−21.6
Coil (C_0)	11.1	36.8	1260.0	58.4 (0.3)	−77.1 (0.3)	−27.6 (0.1)	−46.3 (0.4)	92 (6)	−186.8 (0.6)	−48.9
Coil ($\mathbf{Q} = \mathbf{0}$)					−30.9 (0.4)	−24.7 (0.03)	2.8 (0.5)	50 (8)	−80.8 (0.6)	−33.6
Coil (C_7)	10.6	34.0	1249.0	58.3 (0.5)	−83.4 (0.3)	−28.3 (0.4)	−53.4 (0.8)	98 (4)	−201.4 (0.6)	−57.4

FIG. 2. Hydration free energies of the helix and $\{C_0, \ldots, C_9\}$ coil states. The horizontal axis has no meaning and is used solely to differentiate multiple coil states with similar μ^{ex} values. The radius of the symbol is equal to twice the standard error of the mean (2σ). The C_0 state is the smallest circle in the collection of unfolded states; the standard error is about half compared to the other estimates because we had 4 times more data for C_0 (Sec. S.I).

already noted above, till recently very few all-atom calculations of the hydration free energy of polypeptides with around 10 residues appears to have been reported. Helms and coworkers[18] have studied blocked-(Ala)$_n$ helix hydration with n up to 9. Extrapolating their results suggests a value of about −40 kcal/mol for the deca-alanine helix, consistent with the quasichemical results (Table I). Kokubo et al.[19], computed the van der Waals and electrostatic contribution to the hydration of a deca-alanine helix and for several coil states. Our hydration free energies based on their structures are in fair agreement (data not shown), and the agreement becomes excellent with more extensive sampling in the calculation of the van der Waals contribution [43]. Besides these, the regularization approach has been tested in studies on water[44, 45], ions[46], and simple peptides [11], and the protein cytochrome C[3]. Further, the quasichemical framework has also been thoroughly documented [29–31].

Turning to the results of this study, Table I indicates that at least for the coil states considered here, the packing is somewhat insensitive to the peptide structure. At the scale of the cavities, we expect the packing contributions to scale with surface area [47–49] and the data in the table conforms to this expectation. The free energy per unit SASA is similar to values reported for hard-spheres using scaled-particle theory [49, 50] and explicit all-atom simulations [51–53], but is about a factor of ten larger than the surface energy parameter often used in continuum surface-area based model of hydrophobic hydration [54]. Since the solvent-accessible surface areas of C_0 and C_7 are not very different, despite overall differences in the structure, we expect the packing contribution to be similar for these states, as found in simulations. The packing contribution favors the helix state by about −14 kcal/mol: as expected, hydrophobic hydration favors the compact state of the protein.

The chemistry contribution, however, favors the coil states by between −31.8 kcal/mol for the C_7 state and −25.5 kcal/mol for the C_0 state. Thus, the local protein-solvent interaction outweighs the packing contribution by between 11 kcal/mol to 18 kcal/mol in favoring the coil state. Comparing the chemistry contribution for the peptide and its $\mathbf{Q} = \mathbf{0}$ analog (Table I) shows that the favorable local protein solvent interactions arise primarily from favorable peptide backbone-water interactions, the role of the methyl groups in electrostatic interactions being comparatively negligible. This observation is directly confirmed from the analysis of the solute-solvent interaction contribution to the enthalpy (Table I).

Table I also shows that the entropy of hydration is negative, but here it arises due to attractive solute-solvent interactions. Anticipating our forthcoming article on temperature effects, we note that the entropy calculated using Eq. 5 is in agreement within statistical uncertainties with s^{ex} calculated from the temperature derivative of μ^{ex}, as is expected for thermodynamic consistency.

Results for the $\mathbf{Q} = \mathbf{0}$ analog suggests that attractive solute-water dispersion interactions alone can inhibit chain collapse brought about by packing effects (Table I). The enthalpy of hydration is large negative, despite the positive contribution from solvent reorganization (Table I), emphasizing the role of attractive dispersion interactions between the $\mathbf{Q} = \mathbf{0}$ peptide and solvent. These

observations emphasize that care is needed in assuming the relevance of the poor solubility of nonpolar solutes in rationalizing the collapse of a polypeptide. Our observation that hydration does not explain the collapse of a nonpolar chain is consistent with the observation of similar behavior in alkanes (cf. Ref. 55 and the reanalysis of data in Ref. 56 presented therein). Interestingly, attractive solute-water interactions also oppose the pairing of the prototypical hydrophobe methane [33]; emphasizing the importance of solute-water attractions, this effects is more pronounced for larger alkanes [57].

Protein-solvent long-range interactions (cf. Fig. 1) contribute a substantial fraction — between 50% to 82% — of the net hydration free energy of the peptide. About 90% of the (favorable) long-range contribution for the C_0 and C_7 states arises from van der Waals interactions, while for the helix it is about 60%. As can be expected, the electrostatic contribution for the helix is higher because of the high dipole moment of the helix. (The long-range contributions are non-negligible because the number of solvent molecules in the outer-shell domain is not small, although individual water-solute interaction energies are typically small.) The long-range electrostatic and van der Waals contributions balance in the coil-to-helix transition resulting in a net free energy change of about -4.0 kcal/mol in favor of the helix (Table I).

The above analysis shows that the hydrophilic contributions to hydration outweigh the hydrophobic driving force and favor the *unfolded* state of the peptide. Experiments show that a coil-to-helix transition in a predominantly alanine-based peptide can occur for a polypeptide chain comprising as few as 13 residues [58]. We expect the role of hydration in disfavoring the coil-to-helix transition to hold for this slightly larger chain, provided the coil states are such that the backbone remains accessible to solvent. This then suggests that the experimentally observed coil-to-helix transition must be driven by changes in protein intra-molecular interactions, an inference that is in consonance with the suggestion that enthalpy changes driving helix formation [23], albeit in longer chains. Results in Table II support this expectation.

Table II shows that the favorable hydration of the backbone is lost in coil-to-helix transition ($\Delta E_{bb} > 0$) and this is larger than the change in enthalpy of solvent reorganization ($\Delta E_{reorg} < 0$) which favors the more compact helical state. It is clear that a favorable change in the internal energy ($\Delta E_{int} < 0$) is necessary to obtain a favorable change in the net enthalpy $\Delta h^{\mathrm{ex}} < 0$.

On a per-residue basis the net change in enthalpy in the coil-to-helix transition is estimated to be between -2.4 ± 0.8 kcal/mol/residue (C_0) and -2.8 ± 0.6 kcal/mol/residue (C_7), with statistical uncertainty reported at the 2σ level. A direct comparison of our estimated enthalpy change per residue with experiments is hampered by (a) the lack of a rigorous conformational averaging in our calculations, (b) the short length of our polypeptide, and (c) the fact that in experiments

TABLE II. Components of the enthalpy change in the coil-to-helix transition for the coil states noted below. The change in excess enthalpy (Δh^{ex}) reveals the role of hydration and the change in internal energy (ΔE_{int}) the role of intra-molecular interactions. Δh^{ex}, is further divided into ΔE_{reorg}, the change in the water reorganization contribution, and ΔE_{sw}, the peptide water interaction contribution. The latter is subdivided into contributions from the backbone-water, ΔE_{bb}, and sidechain-water, ΔE_{sc}, interactions. Standard error of the mean is given in parenthesis at the 1σ level. All values are in kcal/mol.

Coil	ΔE_{reorg}	ΔE_{bb}	ΔE_{sc}	ΔE_{int}	$\Delta h_{\mathrm{total}}$
C_0	$-17.0(8.0)$	$34.0(0.4)$	$1.9(0.2)$	-43.3	$-24.0(8.0)$
C_7	$-23.0(6.0)$	$51.1(0.2)$	$-0.6(0.2)$	-55.0	$-28.0(6.0)$

there are residues besides alanine to aid in solubilizing the peptide. Experiments on predominantly alanine peptide (with about 50 residues) [23] suggests a value of -1.0 kcal/mol/residue. Theory suggests that coil states that are less well-hydrated, similar to the C_0 state, will dominate the net hydration thermodynamics (cf. Ref. 59 and 60) for a corresponding result for ions). These are also the states that appear to have an enthalpy change closer to the experimental result (within statistical uncertainties of the calculation). Forcefield bias can be an issue [61], but using the recently re-optimized variant (C36 [62]) of the forcefield changes our results by only an additional 5%.

B. Helix-helix complexation

We next consider the free energy of helix association or the potential of mean force between two helices. Fig. 3 shows that hydration opposes the complexation of helices in the antiparallel orientation. Given that the repulsion starts at a considerable interhelical distance, our choice of helical registration is probably a minor effect. Our results suggest that hydration will oppose formation of the helix-turn-helix motif. Interestingly, the direct intramolecular contributions (ΔU) can outweigh the hydration effects to drive complexation. In the antiparallel arrangement favorable ΔU drives complexation, whereas for the parallel arrangement, unfavorable ΔU tempers the favorable hydration effects. Reminiscent of protein folding free energies, the net free energy of complexation is roughly -2 kcal/mol, a small magnitude relative to the large competing hydration and inter-molecular interaction effects.

Figure 4 shows that primitive hydrophobic effects do drive helix-helix complexation and, in contrast to the coil-to-helix transition, they do outweigh the local chemistry contributions. Thus packing (hydrophobic) effects do become important at larger length-scales. However, the long-range hydration interactions play an entirely nontrivial role in the complexation process, despite the peptides being net neutral; these effects can easily out-
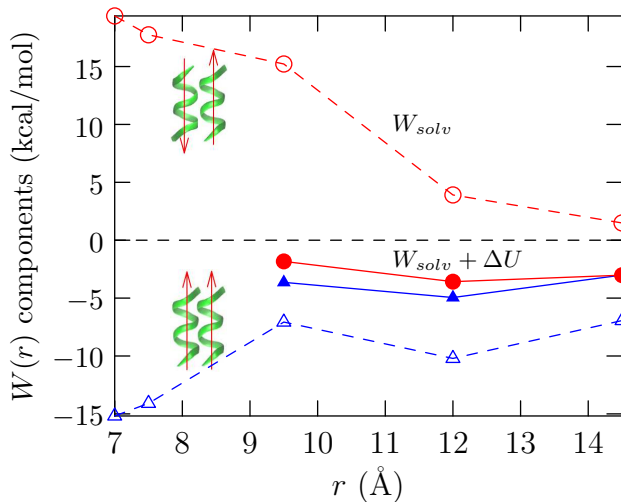
FIG. 3. Components of the potential of mean force in bringing two helices together. The helices are shown in green and the red arrows indicate their mutual orientation. W_{solv} is the solvent contribution (open symbols), and $W_{solv} + \Delta U$ (Eq. 3) is the net PMF (filled symbols). For $r \lesssim 8$ Å, there is steric overlap between the helices and ΔU rises rather sharply. Data including these values of ΔU are thus not shown.

weigh the hydrophobic contribution. The antiparallel arrangement of the helices is strongly disfavored by loss of favorable solute-solvent interactions in hydration, while the inhibition is more modest for the parallel arrangement of helices.

Fig. 5 shows that the orientation dependence of the long-range contribution arises solely due to electrostatic interactions. While van der Waals interactions between the solute and solvent (outside in the first hydration shell) do not discriminate between the two orientations, its magnitude is non-negligible on the scale of the helix complexation free energy. The parallel arrangement of the helices is favored because of the synergistic effect of the helix macro-dipole; the loss of hydration of the helix dipole also explains the unfavorable contribution for the anti-parallel arrangement.

Analysis of enthalpic and entropic effects in pairing shows that for both parallel and antiparallel configurations entropic effects favor complexation (Table III), but enthalpic effects do not. However, the characteristics of the change in reorganization and interaction components emphasizes the need for caution in interpreting the entropic driving force in terms of changes in water structure. For the antiparallel configuration, the change in the water reorganization energy favors helix pairing, but its effect is negligible for pairing of parallel helices. For both orientations, the loss of hydrophilic backbone-solvent and sidechain-solvent interactions inhibits helix association. Emphasizing the importance of electrostatic interactions between the backbone and the solvent, the backbone-solvent contributions are sensitive to the orientation of the helices, but the sidechain-solvent contributions are

essentially of similar magnitude.

IV. CONCLUSIONS

We have considered the hydration contributions to both secondary and tertiary protein structure formation by considering the change of a solvent-exposed coil to helix and the association of two such helices in a model decaalanine peptide. The latter idealized model does ignore the role the loops connecting the helices play in the pairing, but it is a reasonable starting point to understand solvent effects in tertiary structure formation.

The association of the cavities, the prototypical hydrophobic interaction, tends to favor the compact state of the polypetide and favor both the coil-to-helix transition and helix-helix complexation. But in the coil-to-helix transition, hydrophilic effects (protein-water attractive interactions) easily overwhelm the hydrophobic contribution and favor unfolding of the peptide. Even for a discharged peptide, essentially a nonpolar chain, attractive solute-solvent dispersion interactions suffice to favor the unfolded state.

In the pairing of helices, hydrophobic interactions out weigh the short-range peptide-water interactions in favoring helix complexation. This occurs at a larger length-scale than the coil-to-helix transition of a single peptide. However, long-range protein-solvent attractive interactions, especially for the antiparallel arrangement of helices, outweighs the net effect of the packing plus short-range attraction contributions to favor the disassembly of the helices.

In both coil-to-helix transition and helix-pairing, the predominant hydrophilic effects (in our model system) arise from the interaction of the backbone with water. This observed importance of the backbone appears to be consistent with recent studies that encourage a re-appreciation of the role of the backbone in protein folding (for example, see 4–8, 64, and 65).

We find in both the coil-to-helix transition and in the pairing of the helices in the antiparallel orientation, changes in the intra-molecular energy of the protein are essential in shifting the balance to the folded state. The limitations of the models and forcefield notwithstanding, our study suggests that in protein folding hydrophilic effects and protein intra-molecular interactions are as important as, if not more important than, hydrophobic effects.
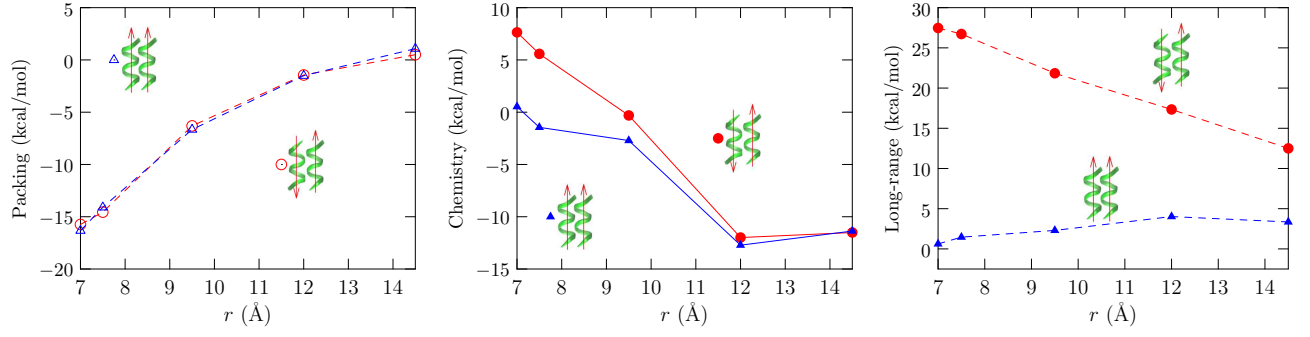
FIG. 4. The SE packing, revised chemistry, and long-range (Eq. 2) contributions to the free energy of helix-helix complexation. The data is presented relative to two helices infinitely apart. At contact ($r \approx 9.5$ Å) the packing effects outweigh the local chemistry effects, but this trend can easily reversed by long-range interactions.

TABLE III. Enthalpic (h^{ex}) and entropic contributions ($T\Delta s^{\text{ex}}$) to the change in the hydration contribution to the potential of mean force, W_{solv}, as the helices are brought from 14.5 Å to 9.5 Å. The change in hydration enthalpy, Δh^{ex}, is further divided into ΔE_{reorg}, the change in the water reorganization contribution, and ΔE_{sw}, the peptide water interaction contribution. The latter is subdivided into contributions from the backbone-water interactions, ΔE_{bb}, and sidechain-water interactions, ΔE_{sc}. $\Delta X = X(9.5 \text{ Å}) - X(14.5 \text{ Å})$, for all X. Standard error of the mean is given in parenthesis at the 1σ level. The value of $Ts^{\text{ex}}(r = 14.5 \text{ Å})$ is explicitly noted to contrast with the sign of ΔE_{reorg}. All units are kcal/mol.

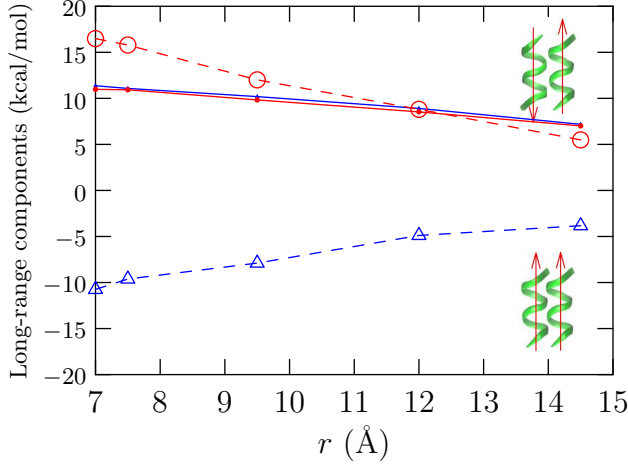| Conformation	ΔE_{bb}	ΔE_{sc}	ΔE_{reorg}	Δh^{ex}	ΔW_{solv}	$T\Delta s^{\text{ex}}$	$Ts^{\text{ex}}(r = 14.5 \text{ Å})$
Antiparallel	24.7 (0.5)	6.2 (0.1)	−7.5 (5.0)	23.4 (5.0)	13.8 (1.2)	9.6	−70.2
Parallel	4.3 (0.4)	4.1 (0.1)	0.8 (5.0)	9.2 (5.0)	−0.1 (1.2)	9.2	−69.3

FIG. 5. The van der Waals (solid line) and electrostatic (open symbols, dashed lines) contributions to the free energy. The long-range contribution to free energy is uniquely decomposable [11, 63] into electrostatic and van der Waals contributions because $P(\varepsilon|\phi) = P(\varepsilon_{vdw}|\phi) \cdot P(\varepsilon_{elec}|\phi)$, where $\varepsilon = \varepsilon_{vdw} + \varepsilon_{elec}$, i.e. the individual binding energy distributions are uncorrelated.

[1] M. M. Lin, D. Shorokhov, and A. H. Zewail, Proc. Natl. Acad. Sc. USA **111**, 14424 (2014).

[2] V. Weber, S. Merchant, and D. Asthagiri, J. Chem. Phys. **135**, 181101 (2011).

[3] V. Weber and D. Asthagiri, J. Chem. Theory Comput. **8**, 3409 (2012).

[4] A. Möglich, K. Joder, and T. Kiefhaber, Proc. Natl. Acad. Sc. USA **103**, 12394 (2006).

[5] S. L. Crick, M. Jayaraman, C. Frieden, R. Wetzel, and R. V. Pappu, Proc. Natl. Acad. Sc. USA **103**, 16764 (2006).

[6] H. T. Tran, A. Mao, and R. V. Pappu, J. Am. Chem. Soc. **130**, 7380 (2008).

[7] C. Y. Hu, G. C. Lynch, H. Kokubo, and B. M. Pettitt, Proteins: Struc. Func. Bioinform. **78**, 695 (2010).

[8] D. P. Teufel, C. M. Johnson, J. K. Lum, and H. Neuweiler, J. Mol. Biol. **409**, 250 (2011).

[9] G. König and S. Boresch, J. Phys. Chem. B **113**, 8967 (2009).

[10] G. König, S. Bruckner, and S. Boresch, Biophys. J. **104**, 453 (2013).

[11] D. S. Tomar, V. Weber, and D. Asthagiri, Biophys. J. **105**, 1482 (2013).

[12] D. S. Tomar, V. Weber, B. M. Pettitt, and D. Asthagiri, J. Phys. Chem. B **118**, 4080 (2014).

[13] C. Tanford, **24**, 1 (1970).

[14] G. I. Makhatadze and P. L. Privalov, Adv. Prot. Chem. **47**, 307 (1995).

[15] W. Kauzmann, Adv. Prot. Chem. **14**, 1 (1959).

[16] C. Tanford, J. Am. Chem. Soc. **84**, 4240 (1962).

[17] K. A. Dill, Biochem. **29**, 7133 (1990).

[18] R. Staritzbichler, W. Gu, and V. Helms, J. Phys. Chem. B **109**, 19000 (2005).

[19] H. Kokubo, C. Y. Hu, and B. M. Pettitt, J. Am. Chem. Soc. **133**, 1849 (2011).

[20] K. A. Dill, S. Bromberg, K. Yue, K. M. Fiebig, D. P. Yee, P. D. Thomas, and H. S. Chan, Prot. Sci. **4**, 561 (1995).

[21] A. S. Yang and B. Honig, J. Mol. Biol. **252**, 351 (1995).

[22] F. Avbelj and L. Fele, J. Mol. Biol. **279**, 665 (1998).

[23] J. M. Scholtz, S. Marqusee, R. L. Baldwin, E. J. York, J. M. Stewart, M. Santoro, and D. W. Bolen, Proc. Natl. Acad. Sc. USA **88**, 2854 (1991).

[24] A. E. García and K. Y. Sanbonmatsu, Proc. Natl. Acad. Sc. USA **99**, 2782 (2002).

[25] S. Gnanakaran and A. E. García, Proteins **59**, 773 (2005).

[26] R. B. Best and G. Hummer, J. Phys. Chem. B **113**, 9004 (2009).

[27] J. M. Richardson and G. I. Makhatadze, J. Mol. Biol. **335**, 1029 (2004).

[28] P. L. Privalov and N. N. Khechinashvili, J. Mol. Biol. **86**, 665 (1974).

[29] M. E. Paulaitis and L. R. Pratt, Adv. Prot. Chem. **62**, 283 (2002).

[30] T. L. Beck, M. E. Paulaitis, and L. R. Pratt, *The potential distribution theorem and models of molecular solutions* (Cambridge University Press, Cambridge, UK, 2006).

[31] L. R. Pratt and D. Asthagiri, in *Free energy calculations: Theory and applications in chemistry and biology*, edited by C. Chipot and A. Pohorille (Springer, Berlin, DE, 2007), vol. 86 of *Springer series in Chemical Physics*, chap. 9, pp. 323–351.

[32] B. Widom, J. Phys. Chem. **86**, 869 (1982).

[33] D. Asthagiri, S. Merchant, and L. R. Pratt, J. Chem. Phys. **128**, 244512 (2008).

[34] L. R. Pratt and A. Pohorille, Proc. Natl. Acad. Sc. USA **89**, 2995 (1992).

[35] L. R. Pratt, Ann. Rev. Phys. Chem. **53**, 409 (2002).

[36] W. Jorgensen, J. Chandrasekhar, . . ., and M. L. Klein, J. Chem. Phys. **79**, 926 (1983).

[37] E. Neria, S. Fischer, and M. Karplus, J. Chem. Phys. **105**, 1902 (1996).

[38] A. D. MacKerell Jr., D. Bashford, . . ., and M. Karplus, J. Phys. Chem. B **102**, 3586 (1998).

[39] A. D. MacKerell Jr., M. Feig, and C. L. Brooks III, J. Comp. Chem. **25**, 1400 (2004).

[40] E. Darve, D. Rodriguez-Gómez, and A. Pohorille, J. Chem. Phys. **128**, 144120 (2008).

[41] J. Hénin, G. Fiorin, C. Chipot, and M. L. Klein, J. Chem. Theory Comput. **6**, 35 (2010).

[42] K. C. Chou, G. M. Maggiora, G. Némethy, and H. A. Scheraga, Proc. Natl. Acad. Sc. USA **85**, 4295 (1988).

[43] H. Kokubo, R. C. Harris, D. Asthagiri, and B. M. Pettitt, J. Phys. Chem. B **117**, 16428 (2013).

[44] V. Weber, S. Merchant, P. D. Dixit, and D. Asthagiri, J. Chem. Phys. **132**, 204509 (2010).

[45] V. Weber and D. Asthagiri, J. Chem. Phys. **133**, 141101 (2010).

[46] Y. Shi and T. L. Beck, J. Chem. Phys. **139**, 044504 (2013).

[47] F. H. Stillinger, J. Soln. Chem. **2**, 141 (1973).

[48] D. Chandler, Nature **437**, 640 (2005).

[49] H. S. Ashbaugh and L. R. Pratt, Rev. Mod. Phys. **78**, 159 (2006).

[50] H. S. Ashbaugh, D. Asthagiri, L. R. Pratt, and S. B. Rempe, Biophys. Chem. **105**, 321 (2003).

[51] P. Varilly, A. J. Patel, and D. Chandler, J. Chem. Phys. **134**, 074109 (2011).

[52] R. C. Harris, J. A. Drake, and B. M. Pettitt, J. Chem. Phys. **141**, 22D525 (2014).

[53] R. C. Harris and B. M. Pettitt, Proc. Natl. Acad. Sc. USA **111**, 14681 (2014).

[54] D. Sitkoff, K. A. Sharp, and B. Honig, J. Phys. Chem. **98**, 1978 (1994).

[55] A. L. Ferguson, P. G. Debenedetti, and A. Z. Panagiotopoulos, J. Phys. Chem. B **113**, 6405 (2009).

[56] M. V. Athawale, G. Goel, T. Ghosh, T. M. Truskett, and S. Garde, Proc. Natl. Acad. Sc. USA **104**, 733 (2007).

[57] D. Ben-Amotz, J. Phys. Chem. Lett. **6**, 1696 (2015).

[58] E. J. Spek, C. A. Olson, Z. S. Shi, and N. R. Kallenbach, J. Am. Chem. Soc. **121**, 5571 (1999).

[59] S. Merchant and D. Asthagiri, J. Chem. Phys. **130**, 195102 (2009).

[60] P. D. Dixit, S. Merchant, and D. Asthagiri, Biophys. J. **96**, 2138 (2009).

[61] R. B. Best, N.-V. Buchete, and G. Hummer, Biophysical J. **95**, L07 (2008).

[62] R. B. Best, X. Zhu, J. Shim, P. E. M. Lopes, J. Mittal, M. Feig, and A. D. MacKerell, Jr., J. Chem. Theory Comput. **8**, 3257 (2012).

[63] S. Utiramerur and M. E. Paulaitis, J. Chem. Phys. **132**, 155102 (2010).

[64] D. W. Bolen and G. D. Rose, Annu. Rev. Biochem. **77**, 339 (2008).

[65] G. D. Rose, P. J. Fleming, J. R. Banavar, and A. Maritan, Proc. Natl. Acad. Sc. USA **103**, 16623 (2006).